

Mapping of Visual Receptive Fields by Tomographic Reconstruction

Gordon Pipa

gpipa@uos.de

Neuroscience Statistics Research Lab, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, U.S.A.; Institute of Cognitive Science, University of Osnabrueck, 49074 Osnabrueck, Germany; and Max-Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany

Zhe Chen

zhechen@mit.edu

Neuroscience Statistics Research Lab, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, U.S.A., and Harvard-MIT Division of Health and Science Technology, Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, U.S.A.

Sergio Neuenschwander

neuenschwand@mpih-frankfurt.mpg.de

Max-Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany

Bruss Lima

lima@mpih-frankfurt.mpg.de

Max-Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany

Emery N. Brown

enb@neurostat.mit.edu

Neuroscience Statistics Research Lab, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, U.S.A., and Harvard-MIT Division of Health and Science Technology, Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, U.S.A.

The moving bar experiment is a classic paradigm for characterizing the receptive field (RF) properties of neurons in primary visual cortex (V1). Current approaches for analyzing neural spiking activity recorded from these experiments do not take into account the point-process nature of these data and the circular geometry of the stimulus presentation. We

*Gordon Pipa and Zhe Chen contributed equally.

Color versions of figures in this article are provided in the online supplement, available at http://www.mitpressjournals.org/doi/suppl/10.1162/NECO_a.00334.

present a novel analysis approach to mapping V1 receptive fields that combines point-process generalized linear models (PPGLM) with tomographic reconstruction computed by filtered-back projection. We use the method to map the RF sizes and orientations of 251 V1 neurons recorded from two macaque monkeys during a moving bar experiment. Our cross-validated goodness-of-fit analyses show that the PPGLM provides a more accurate characterization of spike train data than analyses based on rate functions computed by the methods of spike-triggered averages or first-order Wiener-Volterra kernel. Our analysis leads to a new definition of RF size as the spatial area over which the spiking activity is significantly greater than baseline activity. Our approach yields larger RF sizes and sharper orientation tuning estimates. The tomographic reconstruction paradigm further suggests an efficient approach to choosing the number of directions and the number of trials per direction in designing moving bar experiments. Our results demonstrate that standard tomographic principles for image reconstruction can be adapted to characterize V1 RFs and that two fundamental properties, size and orientation, may be substantially different from what is currently reported.

1 Introduction

The receptive field (RF) of a neuron defines how its spiking activity changes in response to a stimulus. Developing accurate quantitative characterization of the RFs of neurons in primary visual cortex (V1) is critical for understanding how the early visual system represents and transmits information (De Valois, Albrecht, & Thorell, 1982; Jones & Palmer, 1987; de Ruyter van Steveninck & Bialek, 1988; Leventhal, Thompson, Liu, Zhou, & Ault, 1995; Stanley, 2002; Ringach, 2004; Nishimoto, Arai, & Ohzawa, 2005; Chen, Han, Poo, & Dan, 2007). The moving bar experiment is a classic paradigm pioneered by Hubel and Wiesel (1959) for measuring the RFs of V1 neurons. It has been well known that V1 simple cells respond precisely to the location and contrast polarity of features in the visual scene. In the typical moving bar experiment, electrodes are implanted in the animal's V1 area, and a bar is passed in front of the animal's eye (i.e., the visual field) at constant speed and a fixed orientation. The bar is passed systematically at multiple orientations for several trials, and the neurons' spiking activity is recorded across all orientations and all trials. Simple V1 neurons can be well described by a set of linear filters (RFs), which can be analyzed using the principle of reverse correlation (Ringach & Shapley, 2004) or Bayesian inference (Park & Pillow, 2011). The standard approaches to estimating the RF rely on constructing a spatial histogram of the spiking activity or using methods of spike-triggered average (STA) or Wiener-Volterra kernel (WVK) with linear or second-order kernels (Gabbiani, 1996; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997; Paninski, 2003). The STA provides an

unbiased estimate of the linear filter for uncorrelated (or spherically symmetric) stimuli, and WVK provides a minimum-variance estimate in the least-squares sense. The moving bar experiment remains a practical paradigm for mapping visual RFs in awake monkeys, although other alternative visual stimuli have been employed, such as the white noise, grating, and m-sequences (Reid, Victor, & Shapley, 1997).

An examination of the literature found that two important estimation issues have not been considered in standard analyses of V1 moving bar experiments. First, none of these methods uses the fact that neural spiking activity can be accurately represented by point processes (Brown, Kass, & Mitra, 2004; Kass, Ventura, & Brown, 2005; Truccolo, Eden, Fellow, Donoghue, & Brown, 2005; Paninski, 2004). Second, none of these estimation procedures has considered the highly structured geometry of the moving bar experiments (Fiorani, Azzi, & Gattass, 2001). Passing the moving bar at multiple orientations to construct the visual neuron's RF is analogous to tomographic image reconstruction in which energy is passed through an object from different directions. The transmitted energy is captured by arrays of detectors, and a reconstruction algorithm is used to build a tomogram or image of the object's internal structure. For computer tomography, positron emissions tomography, and single photon emission computer tomography, the incident energies are, respectively, X-rays, gamma rays, and electron-positron annihilation (Kak & Slaney, 1988; Deans, 1993; Natterer, 2001). This analogy suggests that tomographic principles for image reconstruction could be used to estimate visual RFs in moving bar experiments. The idea of using the tomographic reconstruction principle to map visual RFs is not new (Sun & Bonds, 1994). However, the spiking activity of V1 neurons was not carefully modeled in that the non-Poissonian firing effect was completely ignored. Another important issue is characterizing the RF size of V1 simple cells. The V1 RF size was traditionally determined based on the minimum-response-field technique (Movshon, Thomson, & Tolhurst, 1978; Pettet & Gilbert, 1992; Crist, Li, & Gilbert, 2001). The minimum-response-field technique assumes the RF profile has a gaussian or difference-of-gaussian shape, and it characterizes the RF extent using a threshold based on the e^{-1} criterion (DeAngelis, Anzai, Ohzawa, & Freeman, 1995) or using the standard deviation to be the size of RF (Nienborg, Bridge, Parker, & Cumming, 2004). However, we argue that this technique lacks statistical justification since, it does not take the uncertainty of the neuronal responses into account; moreover, the gaussian assumption is a poor approximation, and the selection of the threshold is rather ad hoc.

In this article, we combine point-process generalized linear model (PPGLM) methods (Truccolo et al., 2005; McCullagh & Nelder, 1989) with a filtered backprojection algorithm (Kak & Slaney, 1988) to derive a novel analysis approach for mapping the RFs of V1 neurons recorded in a moving bar experiment. Based on our new analysis approach, we propose a statistical criterion (based on the notion of confidence intervals) to define

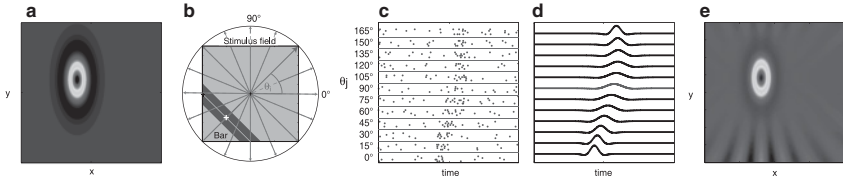


Figure 1: (a) Two-dimensional RF (warm shading represents high firing rate). (b) The bar is moved with a constant velocity with an angle θ_j across the rectangular visual field. (c) Raster plot of a single V1 neuron recorded from one awake monkey with respect to 12 moving bar directions; each row represents recording from one trial in one specific direction. (d) Estimated CIF profiles in 12 directions from panel c. (e) Reconstructed RF using filtered backprojection based on the 12 estimates from panel d. A color version of the figure is provided in the online supplementary material.

the RF size. We use the proposed approach to characterize the RF size and orientation tuning of 251 V1 neurons. Our results show that the orientation preferences of V1 neurons are sharper and the sizes of their RFs are larger than the estimates from conventional methods. Our analysis further suggests an approach to optimal design of moving bar experiments. We emphasize the computational nature and principle of our proposed approach; therefore, an exhaustive comparison with every estimation method in the literature is beyond our scope here. In addition, our proposed approach is specifically aimed at the moving bar experiments; we make no claim about the superiority of moving bar stimuli to other state-of-the-art experimental stimuli paradigms for mapping V1 RFs.

2 Methods

Our approach consists of three important steps: we use a PPGLM to characterize the neuronal responses at each moving bar direction, use the principle of tomography reconstruction to map the visual RF, and characterize the RF size and orientation tuning for every neuron. The first two steps are illustrated in Figure 1.

2.1 Visual Receptive Field Characterization via Radon Transform.

Passing the moving bar at multiple orientations to construct the visual RF is analogous to tomographic image reconstruction in which energy is passed through an object from multiple directions. The neuronal responses to a visual stimulus at different locations of the RF are assumed to follow a linear superposition principle, such that the neural response along a moving bar angle θ is treated as a line integration of the two-dimensional (2D) RF orthogonal to the moving direction.

Mathematically, the Radon transform $\mathcal{R}(f)[\theta]$ specifies the sliced projection of a 2D function $f(x,y)$ along the specific projection angle θ (Kak & Slaney, 1988; Deans, 1993; Natterer, 2001). The V1 RF estimate conditional on the angle θ , characterized as a function of both spatial position (x, y) and orientation α , is denoted as $h(x, y, \alpha|\theta)$. Assuming a factorial property (Rodieck, 1965), the V1 RF is expressed as a product of spatial tuning and orientation tuning: $h(x, y, \alpha|\theta) = f(x, y|\theta)g(\alpha|\theta)$,¹ where $g(\alpha|\theta) \in \mathbb{R}$, $0 \leq \alpha \leq \pi$. It then follows that

$$\mathcal{R}(h)[\theta] = \mathcal{R}(f \cdot g)[\theta] = g(\alpha|\theta) \cdot \mathcal{R}(f)[\theta]. \quad (2.1)$$

That is, the one-dimensional (1D) projection of neural response can be fully characterized by a product of an orientation tuning function and a spatial tuning function jointly defined by the Radon transform. For notation simplicity, the conditional term of the angle θ is made implicit in the functions f and h from now on.

To reconstruct the spatial tuning $f(x,y)$, the inverse Radon transform is applied (Kak & Slaney, 1988),

$$f(x, y) = \mathcal{R}^{-1}\{\mathcal{R}(f)[\theta]\} = \mathcal{F}_2^{-1}\{\mathcal{F}_1\{\mathcal{R}(f)[\theta]\}\}, \quad (2.2)$$

where \mathcal{F}_1 and \mathcal{F}_2 denote the 1D and 2D Fourier transforms, respectively. The second equality follows from the relationship between the Radon transform and Fourier transform. (See appendix A for technical details on continuous Radon transform.) In theory, the projection-slice theorem states that a perfect reconstruction is possible given an infinite number of 1D projections taken at an infinite number of angles (Kak & Slaney, 1988). In practice, the inverse Radon transform is realized by filtered backprojection based on a finite set of projection angles $\{\theta_j\}$, which is implemented numerically in a finite discrete grid.

The spatial tuning function $f(x, y)$ is assumed to have a band-limited spectrum with a bandwidth of b Hz. Reliable reconstruction of $f(x,y)$ based on its line projections $\{\mathcal{R}(f)[\theta_j]\}_{j=1}^J$ requires a minimum sampling frequency q and a minimum number of projections J such that the following two sampling criteria are fulfilled (Natterer, 2001):

$$\begin{cases} J \geq 2b & \text{(angular sampling)} \\ q \geq 2b/\pi & \text{(linear or projection sampling)} \end{cases}. \quad (2.3)$$

¹A more realistic model for the V1 RF is the 2D Gabor filter (Daugman, 1980; Jones & Palmer, 1987), which has coupled or nonseparable spatial and orientation tunings. However, applying the Radon transform to the nonseparable Gabor filter model would make the analysis nontrivial (see equation 2.1).

For $J = 16$ projections evenly spaced on 2π , the bandwidth is upper-bounded: $b \leq 8$ Hz. Arbitrarily small bin size could be chosen to ensure a sufficiently high temporal sampling frequency q .

2.2 Point-Process Generalized Linear Model for Estimating the Radon Transform. In RF mapping, the Radon transform at each moving bar projection direction is unknown and requires estimation based on spike train observations recorded from multiple trials. To do so, single-unit spike trains are treated as point-process observations. Within a time interval $(0, T]$, the spike trains are evenly binned with a bin size of $\Delta = K^{-1}T$ (note that $q = 1/\Delta$). The number K is chosen sufficiently large such that there is at most one spike in any bin, and the point process is modeled by a conditional intensity function (CIF) $\lambda(t_k|H_{t_k})$, for $k = 1, \dots, K$ (Brown, Barbieri, Eden, & Frank, 2003). With a small bin size $\Delta = 1$ ms (i.e., $q = 1000$ Hz), $\lambda(t_k|H_{t_k})\Delta$ approximates the probability of a spiking event within the interval $((k-1)\Delta, k\Delta]$ given the observed spike history H_{t_k} up to time t_k .

For the j th moving bar direction θ_j , we model the log CIF of the spiking activity at time t_k , denoted by $\log \lambda_j(t_k|H_{t_k})$, as the sum of stimulus-evoked and spike history-dependent components (Truccolo et al., 2005),

$$\log \lambda_j(t_k|H_{t_k}) = \sum_{m=1}^M \xi_{m,j}^{(B)} \mathcal{B}_m(t_k) + \sum_{l=1}^I \xi_{l,j}^{(n)} n_{k-l,j}, \quad (2.4)$$

where the first term of the right-hand side of equation 2.4 represents the stimulus-evoked component that is related to the RF characteristic, and the second term represents the spike history-dependent component that is related to the neuron's intrinsic biophysical property. The stimulus-evoked component is modeled using M cubic B-spline functions \mathcal{B}_m , and the spike history-dependent component is modeled using spike count statistics within preceding I history windows. Equation 2.4 essentially defines a GLM with a log link function.

The parameters of PPGLM are estimated using an iterative reweighted least-squares (IRWLS) algorithm (`glmfit` function, Matlab) based on maximum likelihood estimation (Pawitan, 2001; Brown et al., 2003; Paninski, 2004). The order parameters M and I are chosen according to the Akaike's information criterion (AIC) and assessed by a Kolmogorov-Smirnov (KS) test in light of the time-rescaling theorem (Brown, Barbieri, Ventura, Kass, & Frank, 2002; Haslinger, Pipa, & Brown, 2010). Upon completing the maximum likelihood fitting, taking the exponent of both sides of equation 2.4 and evaluating it with the maximum likelihood estimate yields the CIF estimate from the PPGLM (see Figure 1d). In addition, confidence intervals (CIs) for the parameters $\{\xi_{m,j}^{(B)}\}$ and $\{\xi_{l,j}^{(n)}\}$, as well as the CI for $\lambda_j(t_k|H_{t_k})$ at each direction, are analytically derived (Brown et al., 2003).

Due to the circular geometry of the stimulus representation, the neural response along a moving bar angle θ is treated as a line integration of the 2D RF spatial tuning function $f(x,y)$ orthogonal to the moving direction, which is characterized by a Radon transform $\mathcal{R}(f)[\theta]$. At each moving bar direction, the estimated Radon transform is determined only by the stimulus-evoked neural response,

$$\hat{\mathcal{R}}(f)[\theta_j] = \exp \left(\sum_{m=1}^M \hat{\xi}_{m,j}^{(B)} \mathcal{B}_m \right). \quad (2.5)$$

2.3 Modeling Stimulus-Evoked and Spike History-Dependent Components. To model the stimulus-evoked spiking activity of equation 2.4 in each moving bar direction, we use a set of cubic B-spline basis functions $\{\mathcal{B}_m\}_{m=1}^M$. A B-spline is a spline function that has minimal support with respect to a given degree, smoothness, and domain partition. A cubic B-spline is a curve with smoothness order $d = 3$ (i.e., polynomials of degree 3), which is continuous and twice differentiable. The values of control points affect only the shape of equation 2.5 locally in time due to the piecewise definition of \mathcal{B}_m given a sequence of control points $\{0 \leq t_0 \leq t_1 \leq \dots \leq t_{r-1} \leq t_r \leq T\}$ (de Boor, 1978):

$$\mathcal{B}_{m,0}(t) = \begin{cases} 1 & \text{if } t_m \leq t < t_{m+1} \\ 0 & \text{otherwise} \end{cases}, \quad (2.6)$$

$$\mathcal{B}_{m,d}(t) = \begin{cases} \frac{t - t_m}{t_{m+d} - t_m} \mathcal{B}_{m,d-1}(t) + \frac{t_{m+d+1} - t}{t_{m+d+1} - t_{m+1}} \mathcal{B}_{m+1,d-1}(t) & \text{if } t_m \leq t < t_{m+d+1} \\ 0 & \text{otherwise} \end{cases}. \quad (2.7)$$

The control points are placed nonevenly on the 2 s time interval with an adaptive placement strategy, which automatically allocates the control points based on the cumulated peristimulus time histogram (PSTH) of the spike activity (see Figure 2 for illustration).

To model the spike history-dependent component of equation 2.4 at each moving bar direction, we use the spike count within the previous temporal windows to account for the spiking history prior to the current spike's occurrence. The $n_{k-l,j}$ in equation 2.4 is the observed spike count in the past l th temporal window at the j th moving bar direction, and $\xi_{l,j}^{(n)}$ is the associated weight coefficient. A positive (negative) value of $\xi_{l,j}^{(n)}$ implies an excitatory (inhibitory) effect. We let GLM_{hist} and GLM refer, respectively, to the PPGLM with and without the history term, respectively. That is, GLM is a special case of GLM_{hist} with all $\xi_{l,j}^{(n)} = 0$. In the absence of spike

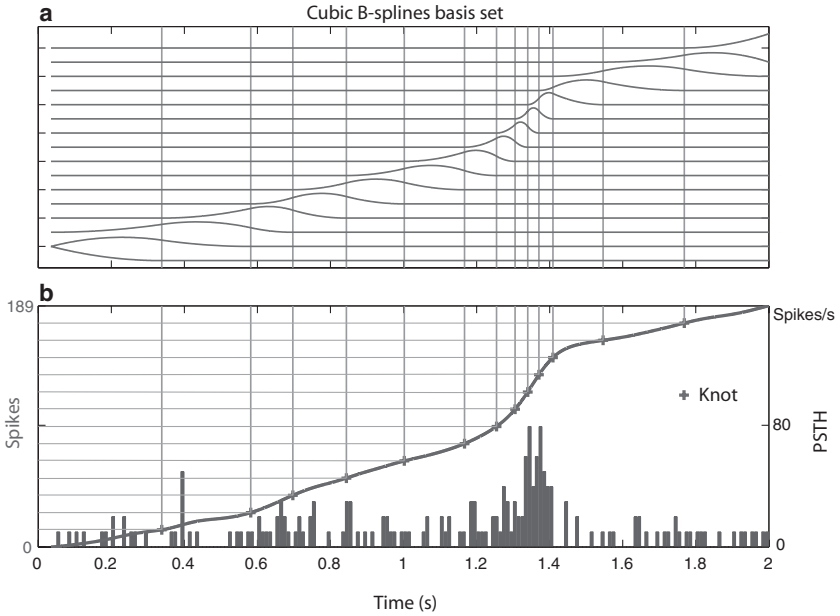


Figure 2: An illustration of adaptive placement of knots for cubic splines. (a) A set of 16 cubic B-splines is placed on the 2 s time interval. The knots are nonuniform in that the distance between two neighboring knots is determined in a way that there are on average the same number of spikes within each time interval. (b) The placement of the knots is determined by the gaussian smoothed cumulative PSTH (right axis, spikes/s).

history dependence, $\lambda_j(t_k|H_{t_k})$ reduces to the standard rate function of an inhomogeneous Poisson process. Therefore, GLM_{hist} is more general since it allows for accommodating non-Possionian neuronal spiking by considering spike history dependence. To fit the GLM_{hist} to the spiking activity for each of the 16 orientations for each of 251 V1 neurons, we use the following five temporal windows: [1–3], [4–6], [7–17], [18–23], [24–35] ms (see Figure 3a for illustration).

2.4 Receptive Field Reconstruction via Filtered Backprojection. Upon estimating the Radon transform in 16 directions, each Radon transform estimate is stored in a 2000×1 vector. In light of equations 2.2 and 2.5, the spatial tuning of the V1 RF is then reconstructed by applying the filtered backprojection algorithm (Kak & Slaney, 1988) to the 16 estimated 1D Radon transforms, followed by a spline interpolation in the finite (1414×1414)

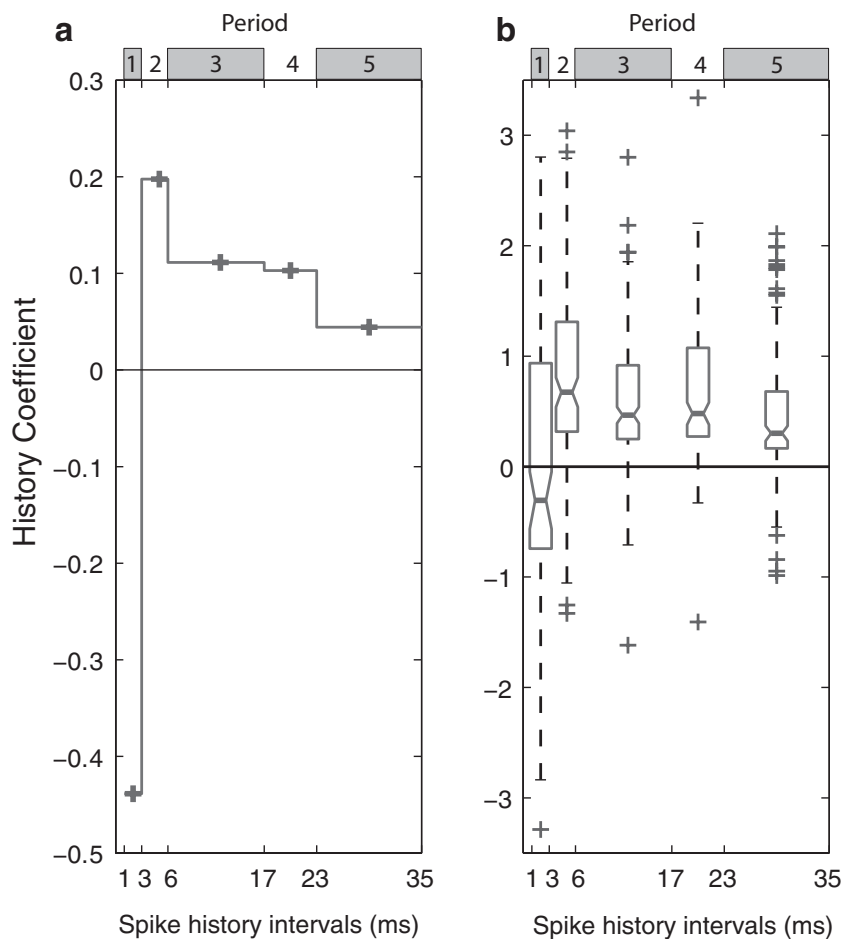


Figure 3: The estimated spiking history-dependent coefficients within five history intervals: [1–3], [4–6], [7–17], [18–23], [24–35] (unit: ms). Five estimated spiking history-dependent coefficients were shown along with their 95% CIs (gray vertical bars). Negative coefficient reflects an inhibitory effect in that interval, indicating the presence of a refractory period. (a) Estimated coefficients from one representative V1 neuron from one moving bar direction. (b) Box plot statistics of the estimated coefficients across all 251 neurons in all 16 directions. The horizontal line within the box indicates the median value. Despite the large variability, the median values of the spike history coefficients have a similar representative profile shown in panel a. A color version of the figure is provided in the online supplementary material.

discrete grid (see Figure 1e).² Specifically, the continuous integration is approximated by a finite sum operation,

$$f(x, y) = \int_0^\pi Q_\theta(t) d\theta \approx \sum_{j=1}^{16} Q_{\theta_j} \quad \text{with } \theta_j = \frac{2(j-1)\pi}{16}, \quad (2.8)$$

where Q_{θ_j} represents a discrete 1D Fourier transform of the filtered version of the estimated Radon transform $\mathcal{R}(f)[\theta_j]$ (Kak & Slaney, 1988),

$$\mathcal{F}_1\{\hat{\mathcal{R}}(f)[\theta_j](t)\} = \sum_{k=-707}^{707} \hat{\mathcal{R}}(f)[\theta_j](t) e^{-i(2\pi w t / 707)}, \quad (2.9)$$

$$Q_{\theta_j}(t) = \text{Re} \left\{ \sum_{m=-707}^{707} \mathcal{F}_1\{\hat{\mathcal{R}}(f)[\theta_j](t)\} |w_m| e^{j(2\pi w_m t)} \right\}, \quad (2.10)$$

where $i = \sqrt{-1}$, and the frequency response $|w|$ in equation 2.10 defines a so-called Ramp filter in the frequency domain (note that $|w|$ is not a square integrable function, so its inverse Fourier transform does not exist in a conventional sense). Therefore, the reconstruction is viewed as a filtered backprojection process: a filtering operation (see equations 2.9 and 2.10) combined with a backprojection (see equation 2.8).

To make the filtered backprojection operation more robust with respect to noise, we use a modified Ramp filter (multiplied by a Hamming window). We specifically construct a symmetric, finite-length discrete filter (with a total length 2048, where 2048 is the next highest power of 2 for 2000) within the Nyquist bandwidth (and therefore cut off the frequency response outside the Nyquist range by setting the response to 0). In addition, we can implement the Ramp filter in the time domain instead of the frequency domain. In so doing, we zero-pad the filter as well as the Radon transform signal, which helps avoid aliasing. We thus replace the circular convolution with a nonperiodic convolution.

Finally, for each discrete point $(\tilde{x}_m, \tilde{y}_n)$ in the 1414×1414 grid, we compute $t_{mn} = \tilde{x}_m \cos \theta_j + \tilde{y}_n \sin \theta_j$ ($1 \leq m, n \leq 1414$), and we use a cubic spline

²The number $1414 = \lfloor 2000/\sqrt{2} \rfloor$ arises from the fact that we treat the moving bar 2 s traveling time (binned as a 2000-dimensional vector) as the diagonal in the rectangular stimulus field (see Figure 1b).

to interpolate $Q_{\theta_j}(t_{mn})$ in the discrete grid from the filtered projection $Q_{\theta_j}(t)$ (see equation 2.10).

Note that there is a typical time delay between the moving bar stimulus and V1 responses in the order of ~ 60 ms. For each direction, we used a fixed latency of 60 ms to shift the neuronal response in time to reflect the delay (otherwise the peak of the estimated RFs would be misidentified). The rationale for estimating the latency and choosing 60 ms is as follows. Given that the moving bar directions are spanned from 0 to 2π (the Radon transform theory requires sampling only from 0 to π by assuming radial symmetry), we can use bars moving in from reverse directions (e.g., 0 and π , $\pi/2$ and $3\pi/2$) to estimate the actual latency. If the latency parameter is correct, two estimated CIFs from opposite directions would be aligned in the peak response; in contrast, if the latency parameter is either too large or too small, the peak responses of two estimated CIFs would deviate. In fact, we have tested other alternative latency values of 50, 55, 65, and 70 ms and found that by using these alternative values, the reconstructed RF is much more elongated, indicating that the latency is either too long or too shorts.³

2.5 Estimating Receptive Field Size and Orientation Tuning. The standard minimum-response-field technique (Movshon et al., 1978; Pettet & Gilbert, 1992) to define the RF extent is based on a gaussian shape assumption of the RF profile. However, the gaussian assumption is not fully justified, and this ad hoc technique lacks statistical justification. Because of the asymmetry of the V1 RF, we compute the RF size from the projected temporal profile at each direction, using two criteria defined below. With the standard e^{-1} criterion (criterion 1; DeAngelis et al., 1995), the RF size is defined by the extent (in time and then converted to degree) where the neuronal response exceeds a threshold (see Figure 4); and the threshold value is defined as e^{-1} of the dynamic range above the baseline firing rate, $\lambda_{j,0} + e^{-1}(\hat{\lambda}_{j,\max} - \lambda_{j,0})$, where e is the Euler's number ($e^{-1} \approx 0.37$); $\lambda_{j,0}$ and $\hat{\lambda}_{j,\max}$ define the baseline and the estimated peak firing rates at the j th direction, respectively; and their difference ($\hat{\lambda}_{j,\max} - \lambda_{j,0}$) defines the dynamic range. The e^{-1} criterion is based on the assumption that the RF is of a gaussian or difference-of-gaussian shape. With criterion 2, the RF size is defined by the extent to which the lower 95% CI of the CIF estimate is greater than the baseline (see Figure 4). The user chooses the level of confidence, which is determined by the need for conservativeness based on the available data. Even with the identical confidence level, a small sample size would often induce a large confidence bound (i.e., uncertainty) to the estimate.

³Alternatively, as suggested by one reviewer, one could apply the reverse correlation or STA method to reveal the period between the stimulus onset and the cortical response that is either significantly correlated or not correlated, and based on that to use it as a rough estimate of the latency.

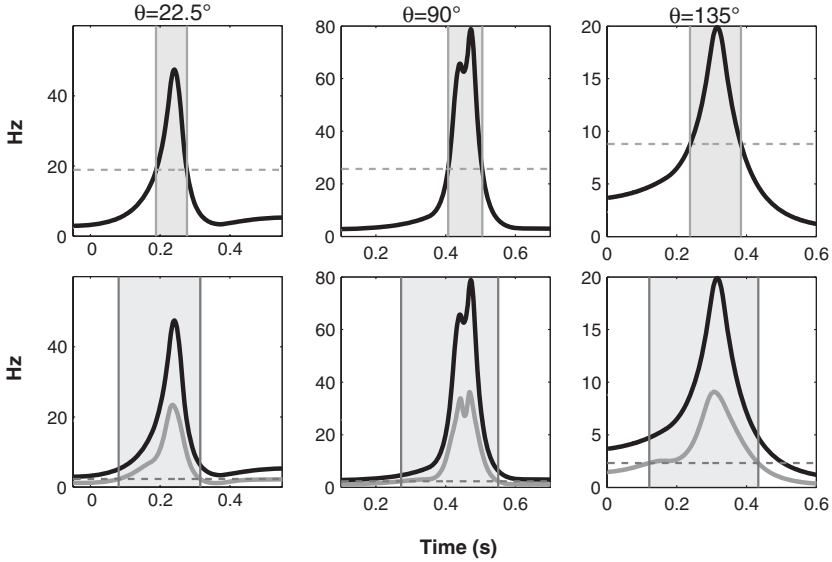


Figure 4: Illustrations of RF size estimation in three moving bar directions for a single V1 neuron based on criterion 1 (top row) and criterion 2 (bottom row), where the black curve indicates the estimated CIF and the shaded area marks the RF range. In the top row panels, the RF size is determined by the e^{-1} criterion, where the horizontal dashed line represents the critical threshold. In the bottom row panels, the RF size is determined by the area where the lower 95% CI of the CIF exceeds the baseline firing rate (horizontal dashed line). A color version of the figure is provided in the online supplementary material.

Specifically, for the j th direction, assuming that the maximum likelihood estimate $\hat{\xi}_j = \{\hat{\xi}_{m,j}^{(B)}, \hat{\xi}_{l,j}^{(n)}\}$ follows a multivariate gaussian distribution $\mathcal{N}(\hat{\xi}_j, \Sigma_j)$ (Pawitan, 2001), then by virtue of equation 2.4, the estimated log-CIF profile $\log \hat{\lambda}_j$ follows a univariate gaussian distribution, $\mathcal{N}(\mu_j, \sigma_j^2)$, and the estimated CIF profile $\hat{\lambda}_j$ follows a log-normal distribution with mean and variance statistics as

$$\mathbb{E}[\hat{\lambda}_j] = e^{\mu_j + \sigma_j^2/2}, \quad (2.11)$$

$$\text{Var}[\hat{\lambda}_j] = (e^{\sigma_j^2} - 1)e^{2\mu_j + \sigma_j^2}. \quad (2.12)$$

The 95% (or 99%) CI roughly corresponds to the twofold (or threefold) standard deviation. In comparison to criterion 1, criterion 2 is statistically meaningful and makes no assumption of the RF shape.

Orientation tuning $g(\alpha|\theta)$ is estimated separately from the spatial tuning $f(x,y)$ at each direction θ_j . The strength of $g(\alpha|\theta)$ in the specific angle θ_j is identified by the estimated peak firing rate $\hat{\lambda}_{j,\max}$ (maximum response during the complete 2 s time interval) and stored in the j th entry of a vector. Upon collecting the values of estimated $\hat{\lambda}_{j,\max}$ from 16 directions, we compute the circular variance (CV) (Mardia, 1972; Ringach, Shapley, & Hawken, 2002),

$$CV = 1 - \left| \frac{\sum_{j=1}^{16} \hat{\lambda}_{j,\max} e^{2i\theta_j}}{\sum_{j=1}^{16} \hat{\lambda}_{j,\max}} \right|, \quad (2.13)$$

and characterize the orientation tuning with an index: $\varepsilon = 1 - CV$ (where $0 \leq \varepsilon \leq 1$). A large value of CV or a low value of ε indicates a low sensitivity to orientation (see Figure 5c). Since the periodicity of orientation is 180° , but the moving bar direction is cyclic over 360° , the actual directions of the bar are used to calculate the directional preferences of each neuron. In all graphic visualizations presented later (see Figures 5b and 7), instead of visualizing $h(x, y, \alpha) = f(x, y)g(\alpha)$, we plot a rescaled version of $f(x, y)$ by assuming an isotropic orientation in all directions.

3 Results

3.1 Data. The spiking activity from 251 V1 neurons was recorded in two awake macaque monkeys during the moving bar experiment (see Figure 1c for the raster plot of one V1 neuron), in which a moving light bar (luminance: $50 \text{ cd}\cdot\text{m}^{-2}$, bar width: 0.2°) stimulus was used in 16 different orientations separated evenly by 22.5° in nine trials per direction presented in random order. Each trial lasted 2 seconds. Details of experimental setup and recordings are described in appendix C.

3.2 Encoding Analysis and Receptive Field Mapping. In the encoding analysis, we estimate the 1D Radon transform of each of the 16 different orientations as the CIF of a PPGLM fit to the spiking activity. We compare the reconstructed RF from two PPGLMs—one using spike history (GLM_{hist}) and one not using spike history (GLM)—with estimates computed from standard methods of STA and WVK (using only linear kernels). The details of the STA and WVK methods are given in appendix B.

Figure 5 shows a typical comparison of four estimation methods. All provide rate function estimates in each of the 16 directions from the nine trials of spiking activity (see Figure 5a). In Figure 5a, the estimated CIFs are plotted against the raw spike rasters. The CIFs from the GLM_{hist} and the GLM are smoother than those from the STA and WVK because the former two used smoothing cubic splines as basis functions. The WVK and STA rate function estimates are noisier because neither considers the point-process

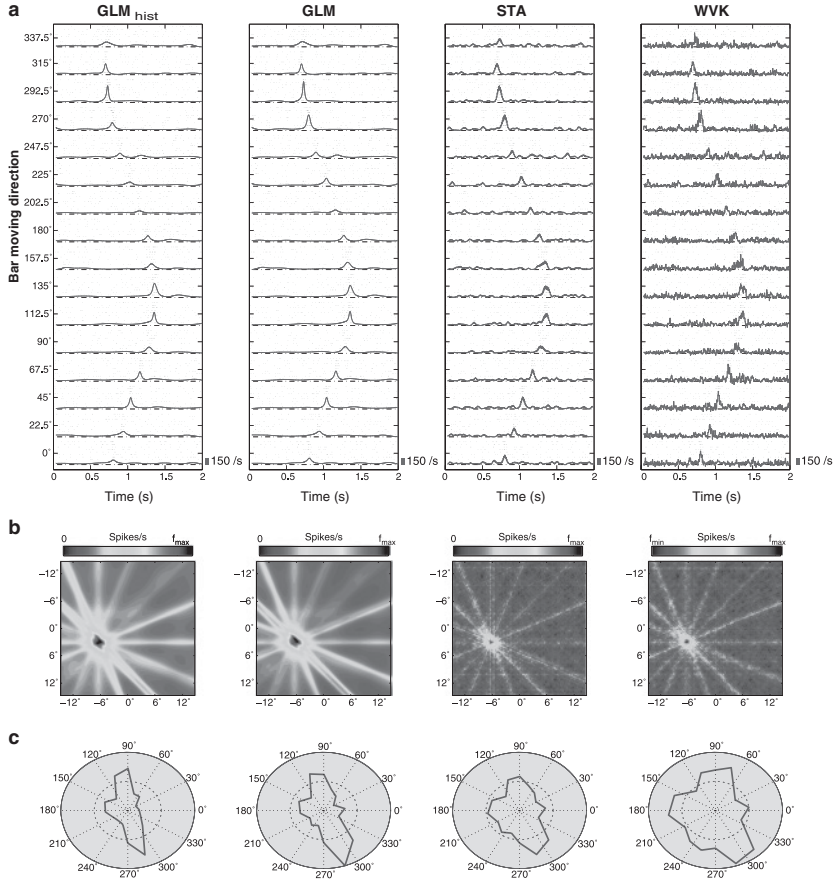


Figure 5: Comparison of four spike encoding approaches (columns 1–4: GLM_{hist}, GLM, STA, WVK) in estimating an excitatory RF of V1 neuron. (a) Raster plot of the single cell response to 16 moving bar directions, with each direction containing 9 experimental trials. Each panel shows the spike train data with the estimated stimulus response (spikes/s) in each direction. (b) The estimated 2D RF is shown in terms of the visual angle (in degrees) relative to the fixation point. Response strength is coded for spatial tuning. In this example, the estimated peak rates f_{max} from four paradigms are 70, 90, 90, and 120 spikes/s, respectively. The mean RF size estimate averaged across all 16 directions is GLM_{hist}: 1.46°; GLM, 1.36°; STA, 1.39°; WVK: 1.18°. (c) Polar plot for the RF orientation tuning estimate, characterized by an ε -index. The ε -estimates from four approaches are, respectively, 0.24, 0.23, 0.13, and 0.13. The GLM_{hist} estimate for the RF size is larger and has a sharper orientation tuning than the STA and WVK estimates. A color version of the figure is provided in the online supplementary material.

Table 1: Goodness-of-Fit Assessments Summary for the Fits of the GLM_{hist} , GLM, STA, and WVK Models for 251 V1 Neurons.

	GLM_{hist}	GLM	STA	WVK
Complete data (9 trials per direction)	238 (94.9%)	171 (68.1%)	154 (61.4%)	129 (51.4%)
Training data (6 trials per direction)	243 (96.8%)	183 (72.9%)	162 (64.5%)	128 (51.0%)
Validation data (3 trials per direction)	242 (96.4%)	179 (71.3%)	112 (44.6%)	61 (24.3%)

Note: The table entries are the number (and percentage) of the neurons whose KS plots are within the 95% CIs, suggesting a close agreement between the model and experimental data.

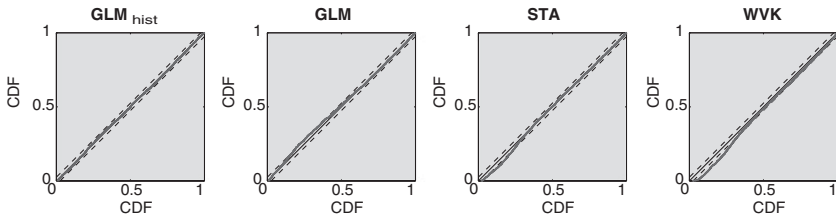


Figure 6: The KS plot comparison between four spike encoding methods in modeling the V1 neuronal spiking data used in Figure 4. The horizontal and vertical axes of these plots are the time-rescaled empirical cumulative distribution function (cdf) and the theoretically uniform cdf, respectively. Here, only the KS statistic from GLM_{hist} passes the test, in which the curve falls completely within the 95% CI.

nature of the spiking activity. The WVK is a least-squares estimate whereas the STA is the simple average of the spike counts over the nine trials. Furthermore, the RF spatial tunings of the GLM_{hist} and the GLM are sharper than those of the STA and the WVK (see Figure 5b). The spatial RF's peak amplitude of the GLM_{hist} is typically less than that of the GLM because the latter does not consider spike history. Note that all spatial tuning estimates have stripe-like “artifacts” that are clearly related to the discrete sampling of the bars at the 16 orientations. In addition, the GLM_{hist} gives sharper orientation tuning measured as a greater ε -index value than the GLM, STA, and WVK (see Figure 5c).

Of the four models considered, the GLM_{hist} gives the most accurate description of the spiking activity across all of the 251 V1 neurons in terms of cross-validation KS goodness-of-fit assessments (see Table 1 and Figure 6 for one example). In the validation data, the GLM_{hist} fits meet the KS goodness-of-fit criterion for 96.4% of the neurons compared with 71.3% of the GLM, 44.6% of the STA and 24.3% of the WVK fits. Therefore, in the subsequent analyses, we compare the GLM_{hist} results with the conventional STA and WVK methods.

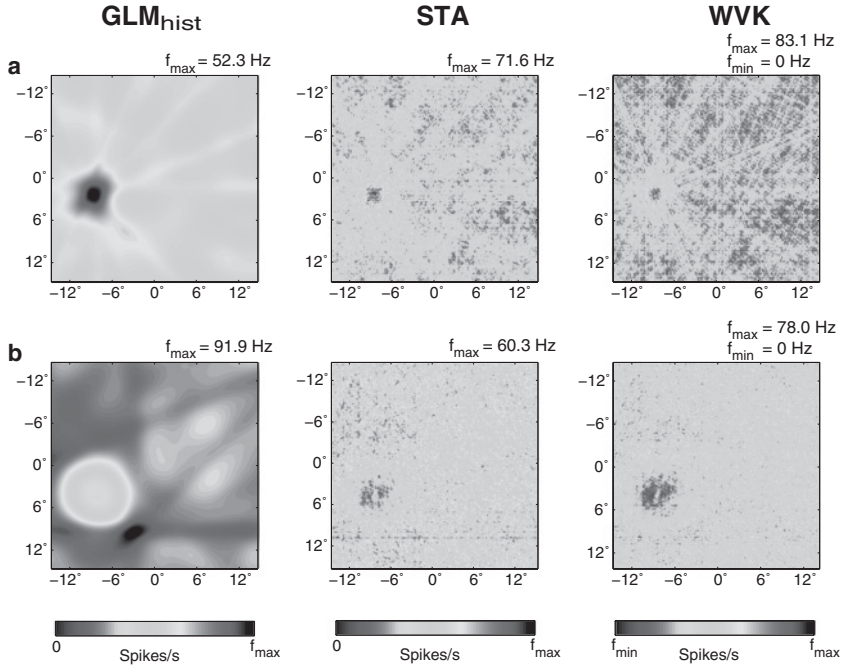


Figure 7: Comparison of three methods in the RF mappings of two different types of V1 simple cells: an excitatory cell that has an above-baseline response to the stimulus (a) and an inhibitory or suppressive cell that has a below-baseline response to the stimulus (b). The RF is shown as a function of the visual angle relative to the fixation point. The neural response of the spatial tuning function is represented in grayscale. At the top right corner of each panel, the estimated maximum (or minimum) rate is marked. A color version of the figure is provided in the online supplementary material.

V1 neurons have both excitatory and inhibitory RFs. Two representative examples of spatial RF mapping are shown in Figure 7. In the encoding analysis, GLM_{hist} consistently obtains a smoother and sharper spatial RF estimate (higher contrast and spatially well defined).

3.3 Receptive Field Size, Orientation Tuning, and Bandwidth. Our PPGLM framework also provides a new criterion for assessing RF size based on comparing the baseline firing activity with the lower 95% CI of the CIF. The logic behind this new definition is that activity significantly greater than baseline can be considered the stimulus-evoked response. Using the new criterion, we find that the estimate of RF size is asymmetric and larger than that determined by the e^{-1} criterion. As illustrated in Figure 4 for one V1

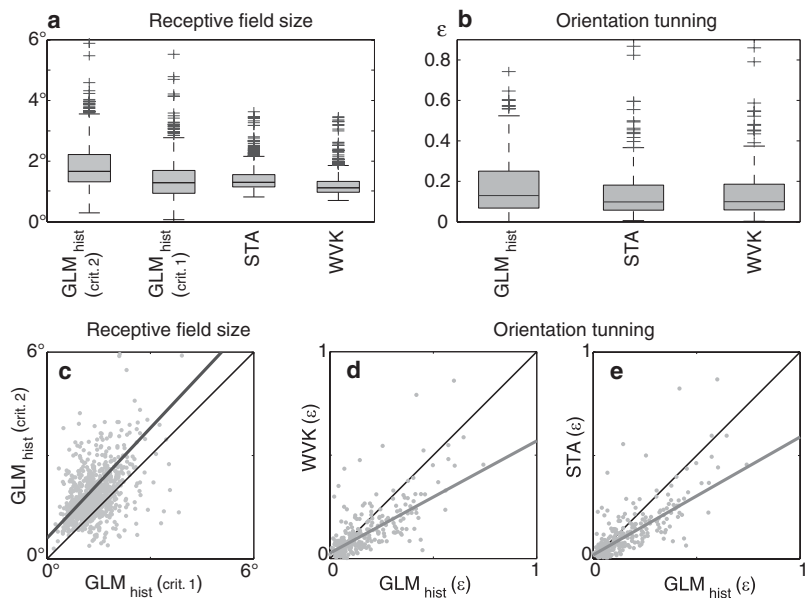


Figure 8: Summary statistics of RF size and orientation tuning estimates for 251 V1 neurons. (a, b) Box plots comparison of three spike encoding methods in estimating the RF size and orientation tuning. On each box, the center line marks the median, the edges of the box are the 25% and 75% quantiles, and the whiskers indicate 1.5 times the interquartile range from the edges of the box. Outliers are marked by the symbol +. The RF size estimates from GLM_{hist} (criterion 2) differ significantly from three other estimates ($n = 4,016$ orientations, $P < 0.01$, Kruskal-Wallis one-way ANOVA median test, and a follow-up post hoc pairwise median test), and the other three show no difference among themselves ($P > 0.1$). The orientation tuning estimates differ significantly between GLM_{hist} and STA and WVK ($n = 251$ cells, $P < 0.01$), and the estimates from STA and WVK are not statistically different ($P > 0.1$). (c) Scatter plot of the RF size estimates from GLM_{hist} using two criteria. The two estimates appear correlated, but the estimates using criterion 2 are larger (least-squares fit: $RF_{GLM, crit. 2} = 1.066 \times RF_{GLM, crit. 1} + 0.599$, shown in the thick gray line). (d, e) Scatter plots of the ϵ -estimates between GLM_{hist} and the STA and WVK methods. The slope and its 95% CI of a linear regression fit are calculated: slope = 0.543 ± 0.021 for WVK versus GLM_{hist}; slope = 0.573 ± 0.023 for STA versus GLM_{hist}. A color version of the figure is provided in the online supplementary material.

neuron, the RF size estimates from three selected moving bar directions are smaller using the old criterion than using the new one. This is consistently observed across the majority of the orientations (80.9% among all 16×251 orientations; see Figure 8c) of the V1 neurons.

We compare the population RF size and orientation estimates of three methods: GLM_{hist} , STA, and WVK (see Table 1). For GLM_{hist} , we consider two criteria for defining RF size: criterion 1, the standard e^{-1} criterion, and criterion 2, defined by the 95% CI of the estimated CIF. For STA and WVK, we consider only the standard e^{-1} criterion. In principle, we can also use the 95% CI criterion for the WVK (i.e., roughly two standard deviations of the estimate; see the appendix); however, the least-squares method may produce a negative CIF estimate at some time points (see the last column of Figure 5a for an illustration), such that its lower 95% CI is completely within the negative region. In this case, it is not meaningful at all to use the 95% CI criterion for the WVK method.

As seen from Figure 8a, based on criterion 1, the median RF size estimates from GLM_{hist} , STA, and WVK are in close agreement (pairwise median tests between STA and WVK, $n = 4016$ orientations, $P > 0.05$). Using criterion 2, GLM_{hist} has a larger median RF size. The RF size estimates from the GLM_{hist} using criterion 2 are linearly related to, but significantly larger than, the estimate from GLM_{hist} using criterion 1 (see Figure 8c; linear regression, $\text{RF}_{\text{GLM,crit 2}} = 1.066 \times \text{RF}_{\text{GLM,crit 1}} + 0.599$), suggesting a clear offset between these estimates. In the analysis of the orientation tuning (see Figure 8b), the ε -estimate from GLM_{hist} is significantly different from STA and WVK ($n = 251$ cells, $P < 0.01$, Kruskal-Wallis one-way ANOVA median test). No significant difference is found between STA and WVK estimates. In addition, the ε -estimate from GLM_{hist} is nearly twice the estimates from WVK and STA (see Figures 8d and 8e). The average slope ($\pm 95\%$ CI) for the linear regression of WVK versus GLM_{hist} is 0.543 ± 0.021 , and for STA versus GLM_{hist} it is 0.573 ± 0.023 . These findings suggest that formal statistical modeling of the point-process nature of the neural spiking activity, combined with tomographic reconstruction and statistically based definition of the RF property, leads to substantially different inferences about the RF size and orientation tuning. Note that there is no significant correlation between the estimated RF size and the orientation tuning. We compare the median RF size and the ε -estimate from 16 directions from 251 neurons; the Pearson correlation coefficients for the GLM_{hist} , STA, and WVK methods are -0.1284 , -0.1548 , and -0.1122 , respectively. None of these correlations is statistically significant.

Furthermore, using an approximate formula between the index ε and the V1 RF bandwidth b (Ringach et al., 2002),

$$\varepsilon \approx \left(\frac{\sin(2\pi b)}{2\pi b} \right)^2, \quad (3.1)$$

we estimate the V1 RF bandwidth b from the ε -estimate. A greater ε -estimate (a sharper orientation selectivity) yielded a narrower bandwidth. The median value 0.129 (25% quantile: 0.067, and 75% quantile: 0.250, respectively)

of the ε -estimate among 251 V1 neurons (see Figure 8b) corresponds to a median bandwidth 0.35 Hz (0.39 Hz and 0.30 Hz, respectively).

3.4 Experimental Design. For experiments using moving bar stimuli to estimate visual RF, two design questions arise naturally. What is the minimum number of moving bar directions required to reconstruct the RF with high fidelity? Given a fixed total number of trials, what is the optimal configuration of number of directions and the number of trials per direction?

To answer these experimental design questions, we simulate spiking activity from a V1 neuron with an ON-center, OFF-surround RF (with a shape of the difference of gaussians) having the following spatial tuning function (see Figure 9a),

$$f(x, y) = A_1 \exp\left(-\frac{0.5(x^2 + y^2)}{\sigma^2}\right) - A_2 \exp\left(-\frac{0.5(x^2 + y^2)}{1.2\sigma^2}\right) + A_3, \quad (3.2)$$

with $\sigma^2 = 1/8$. The function $f(x, y)$ is properly shifted and scaled (by three parameters A_1 , A_2 , and A_3) such that the ultimate function $f(x, y) > 0$ and has a peak firing rate of 125 Hz and a baseline firing rate of 25 Hz. The RF $f(x, y)$ is discretized in a 1414×1414 grid and has a bandwidth of 8 Hz/pixel.

From equation 2.13 we simulate 2 seconds of spike trains in each of 128 evenly spaced directions with nine trials per direction. From these data, we reconstruct the RF using 8, 16, 32, 64, and 128 directions and one to nine trials per direction (see Figure 9 for illustrations). Given a number of directions and a number of trials per direction, we fit PPGLMs to the spiking activity to estimate the Radon transforms in each direction and then applied filtered backprojection to the estimated transforms to reconstruct the RF. We evaluate the quality of the reconstruction by computing the mean-squared reconstruction error (MSRE) between the estimated RF and the true value across all 1414×1414 pixels.

As expected, the reconstruction error decreases as the number of directions and the number of trials per direction increases (see Figure 10a). In examining the MSRE with respect to the total number of trials, we observe that, provided the condition for the minimum number of spatial samples is satisfied (at least 16 directions), the error decreases with an increasing number of trials. However, as the total number of trials reaches approximately 200, the MSRE approaches an asymptote of approximately 2×10^{-5} (see Figure 10b). As the sampling theory suggests, if the high-frequency details of the RF are required (i.e., larger bandwidth), then the number of directions is crucial, and oversampling will yield a more accurate reconstruction. Because in moving bar experiments the neural responses are stochastic, a minimal number of trials per direction is always required. We find that to achieve an MSRE of 2×10^{-5} , a feasible design would be to use 32

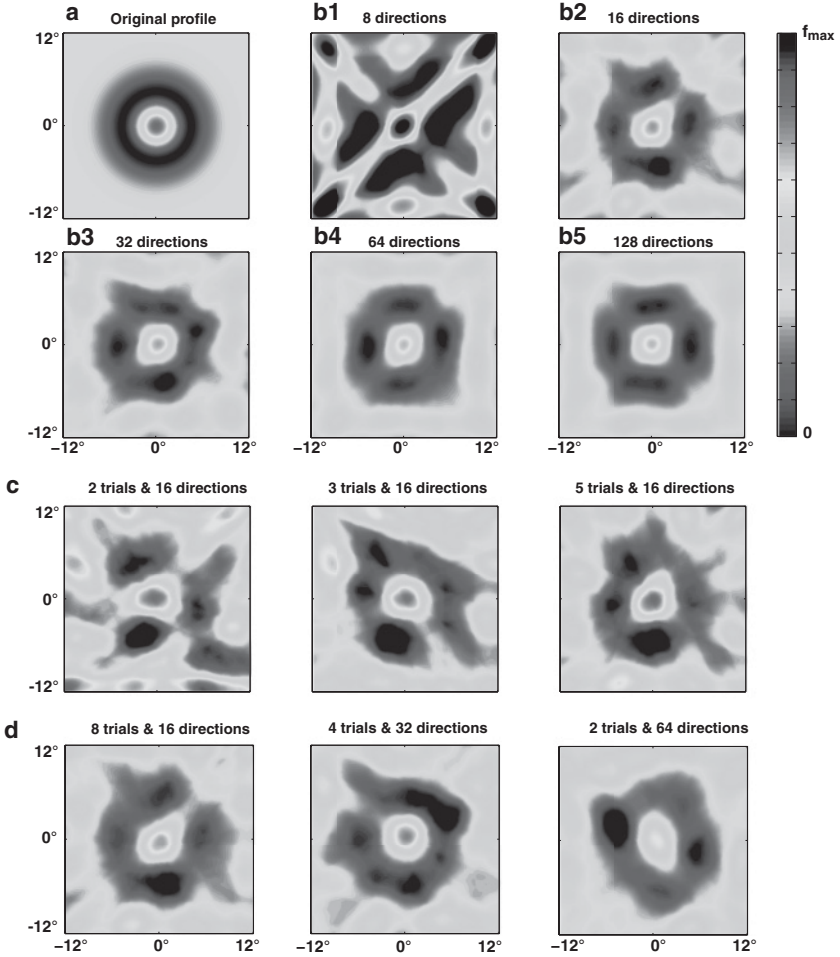


Figure 9: (a) The simulated 2D visual RF, and the reconstructed RF based on tomographic reconstruction using various numbers of projections (b1–b5). The visual RF is simulated with a baseline firing rate of 25 Hz and a peak firing rate of $f_{\max} = 125$ Hz. From the visual RF, inhomogeneous Poisson spike trains are generated along all directions. In each direction, nine trials are simulated. According to the Radon transform sampling criterion, to ensure a reliable signal reconstruction, at least 16 directions are required. In panel b1, due to undersampling effect, the RF reconstruction is very poor. As more and more projections are added, the fidelity of the reconstructed RF improves. (c) Reconstruction comparison using varying numbers {2, 3, 5} of trials in each direction for the same number of 16 directions. (d) Reconstruction comparison using three different configurations with a fixed total number of 128 trials. A color version of the figure is provided in the online supplementary material.

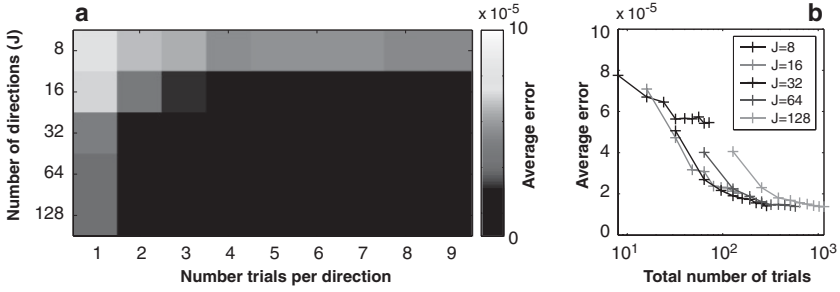


Figure 10: Results from the simulation study using a difference-of-gaussians model. (a) Comparison of the MSRE (normalized by the total number of pixels) with varying number of directions and number of trials per direction. (b) The MSRE curve as a function of the total number of trials for a fixed number of directions J . The MSRE saturates to an asymptotic level of approximately 2×10^{-5} when the total number of trials approaches about 200. A color version of the figure is provided in the online supplementary material.

directions, with at least 6 trials per direction, or 192 trials for the experiment, which suggests a different configuration from the 16 directions and 9 trials per direction used in our experiment.

3.5 Robustness of the Separability Assumption. Thus far, we have assumed that the spatial tuning and orientation tuning of the V1 neurons are separable (see equation 2.1). In reality, V1 neurons are better modeled by a 2D Gabor filter (Daugman, 1980; Jones & Palmer, 1987). In order to test the robustness of our assumption and tomographic reconstruction approach, we further simulate a V1 neuron using a Gabor filter model, which is given by the product of a gaussian-shape envelope and a complex sinusoidal carrier (Movellan, 2008),

$$\begin{aligned}
 f(x, y) = & A \exp \left(- \frac{\pi ((x - x_0) \cos \xi + (y - y_0) \sin \xi)^2}{a^2} \right. \\
 & \left. - \frac{\pi (-(x - x_0) \sin \xi + (y - y_0) \cos \xi)^2}{b^2} \right) \\
 & \times \exp(i2\pi(u_0x + v_0y) + i\phi),
 \end{aligned} \tag{3.3}$$

where (x_0, y_0) denotes the location of the peak of the gaussian envelope; (a^2, b^2) denote the spatial variances along the (x, y) coordinates, respectively; (A, ξ) denote the amplitude and rotation angle of the gaussian envelope, respectively; (u_0, v_0) and ϕ denote the spatial frequencies and phase of the sinusoid, respectively. Taking the 2D Fourier transform of equation 3.3

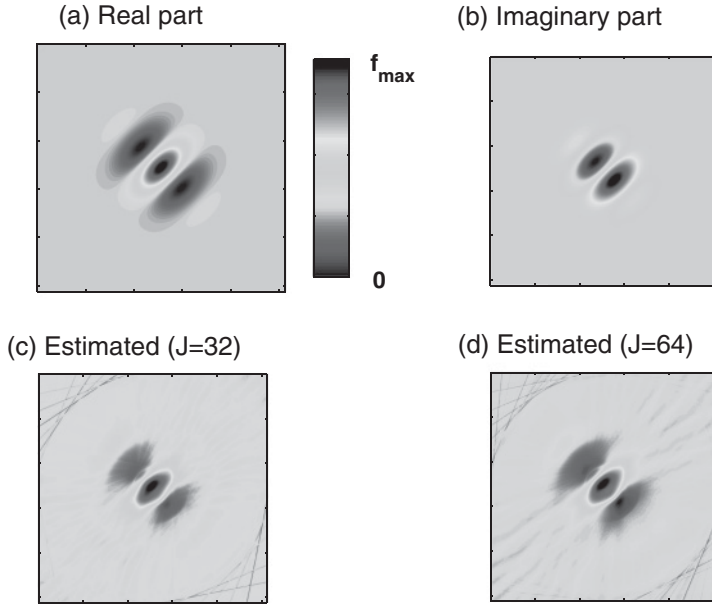


Figure 11: Results from the simulation study using a Gabor filter model. (a) The real part of the complex Gabor filter in spatial domain ($f_{\max} = 80$ Hz). (b) The imaginary part of the complex Gabor filter in spatial domain. (c) The estimated RF profile based on 32 projections. (d) The estimated RF profile based on 64 projections. A color version of the figure is provided in the online supplementary material.

further yields

$$\begin{aligned} \mathcal{F}(u, v) = & \frac{A}{ab} \exp(-i2\pi(x_0(u - u_0) + y_0(v - v_0)) + \phi) \\ & \times \exp\left(-\pi \frac{((u - u_0) \cos \xi + (v - v_0) \sin \xi)^2}{a^2} \right. \\ & \left. - \pi \frac{(-(u - u_0) \sin \xi + (v - v_0) \cos \xi)^2}{b^2} \right). \end{aligned} \quad (3.4)$$

Basically, the Gabor filter is spatially localized and can be viewed as a bandpass filter that favors spatial frequencies within a certain range. Here, we set $x_0 = y_0 = 0$, $a = 50$ pixels, $b = 40$ pixels, $\xi = \pi/4$, $u_0 = v_0 = 1/80$ cycles/pixel, and $\phi = 0$. Let $C = \sqrt{\frac{\log 2}{\pi}}$ and $R = \sqrt{u_0^2 + v_0^2}$, the spatial frequency BW is about $2C/a \approx 0.0188$ pixel, and the orientation BW is about $2 \tan^{-1}(C/Rb) \approx 1.173$ deg (see Movellan, 2008, for BW derivations based on the half-magnitude RF profile). Figures 11a and 11b show the real and

imaginary parts of the Gabor filter in spatial domain on an image of 1024×1024 pixels, respectively.

From the real part of the Gabor filter (which is properly shifted such that it is nonnegative, with a peak firing rate of ~ 80 Hz, and a baseline firing rate of ~ 20 Hz), we simulate 2 seconds of spike trains in 32 (or 64) evenly spaced directions in $[0, 2\pi)$ with four trials per direction.⁴ Using the separability assumption, we estimate the spatial tuning of the visual RF and evaluate the quality of the reconstruction by computing the MSRE between the ground truth (see Figure 11a) and estimated RF profile (in terms of the pixel intensity of two images). The estimated results are shown in Figures 11c and 11d. It appears that despite the use of the separability assumption, the estimated RF profile remains qualitatively and quantitatively similar to the true RF. In these two cases, the MSRE values are 0.0432 and 0.0304, respectively. As seen in Figure 11, although the details of the reconstructed RF images are partly distorted, the underlying Gabor filter structure has been captured, suggesting the robustness of the separability assumption.

4 Discussion

4.1 Tomographic Reconstruction and PPGLM. By combining statistical modeling of neural spike trains using the PPGLM framework with the principles of tomographic reconstruction, we have presented a new two-step approach to estimating the RFs of V1 neurons from spiking activity recorded in an experiment in which a bar was moved across the visual field at multiple orientations. In the first step, a PPGLM is used to estimate the Radon transform as the CIF of the spike trains recorded across multiple trials in each direction. In the second step, under the linear superposition principle, we invert the estimated Radon transforms to reconstruct the V1 RF using filtered backprojection. This approach takes advantage of the point-process nature of the spiking activity and the highly structured geometry of the experiment. This latter step is analogous to image reconstruction in which use of filtered backprojection is standard practice (Kak & Slaney, 1988; Deans, 1993). Our use of tomographic principles to map V1 RFs is a further extension of previous work by Sun and Bonds (1994) in two important aspects: one is characterization of spiking activity via a PPGLM; the other is to propose a statistical criterion for determining the RF size.

Our PPGLM (GLM_{hist}) that models spike history along with stimulus effect gives the most accurate description of the spiking data compared with a PPGLM that models only the stimulus effect (GLM) and the standard STA and WVK algorithms for rate function estimates. There are several

⁴Assuming circular symmetry, this is equivalent to sampling from 16 (or 32) evenly spaced directions in $[0, \pi)$ with eight trials per directions.

reasons that the GLM_{hist} outperforms the STA and WVK. Instead of treating the spiking activity as a point process, the WVK use a linear gaussian model to relate the visual stimuli to neural responses (Paninski, 2003). The WVK parameters are one of the initial guesses commonly used to start the IRWLS algorithm used to compute the PPGLM parameter estimates (Pawitan, 2001). The STA is an approximation to the WVK. Neither of these methods takes into account spike history. The spike history effect reflects the nature of non-Poissonian firing (Shimokawa & Shinomoto, 2009), as evidenced in most V1 neurons (see Figure 3b). Compared to coarse spike count coding (as in STA and WVK), fine temporal coding contains more information of visual stimuli, as seen by its superior decoding performance in real visual system (Jacobs et al., 2009). It is also now appreciated that intrinsic neuronal dynamics including absolute and relative refractory periods are important for predicting current spiking activity (Truccolo et al., 2005; Truccolo, Hochberg, & Donoghue, 2010). Including both the stimulus and spike history in the PPGLM helps explain its superior performance relative to the STA and WVK in terms of the accuracy of the data description and the model generalizability, as evaluated, respectively, in the goodness-of-fit and cross-validation analyses. Since neither STA nor WVK meets the sampling criteria imposed by tomographic reconstruction, they consequently suffer a loss of reconstruction accuracy.

4.2 Receptive Field Size and Uncertainty. The RF of a V1 neuron defines a region of space that is sensitive to a visual stimulus. Such a region can have rather complex dependencies of the stimulus sensitivity on the exact position and type of the stimulus. Despite this complexity, an RF is often experimentally described with rather simple measures, such as the RF size and the orientation specificity. Our analysis, based on the maximum likelihood estimation of the PPGLM, provides a new definition of RF size. This definition is more reasonable than the standard one because it uses an accepted statistical criterion. The spatial extent of the RF is determined by the region over which the lower bound of 95% CI of the CIF is greater than the baseline. In contrast, the e^{-1} criterion assumes a gaussian shape for each RF and does not consider the statistical properties of the data. In particular, it is possible to set the e^{-1} boundary beyond the point where there is a statistically significant increase in the CIF relative to baseline. It is also possible to have statistically significant activity beyond the boundary defined by the e^{-1} criterion. This latter case is what we have observed. In the GLM_{hist} analyses, the RF sizes based on the new criterion are systematically larger compared with the RF sizes estimated by the e^{-1} criterion (see Figure 8c). The GLM_{hist} RFs also have greater orientation specificity than those estimated by either the STA or the WVK (see Figures 8d and 8e).

One has to keep in mind that defining the RF size requires determining a threshold or minimal sensitivity given a certain type of a stimulus. Based on the signal detection theory, it is clear that changing the threshold or

minimal sensitivity results in changes of the RF size. Given the typical shape of the RFs in V1, a stricter criterion that requires a larger sensitivity will typically result in a smaller RF size and vice versa. We expect the same for the threshold criterion and the CI criterion discussed in this article. For both criteria, the precise change of the RF size as a function of the threshold will depend on the exact shape of the spatial sensitivity. In the case of the CI criterion, the threshold itself also depends on the number of data used for fitting the PPGLM, so more data usually lead to smaller CIs and larger RF size. Therefore, the RF size estimate should be seen as not only an intrinsic property of the V1 neurons, but also an induced quantity of uncertainty about the data. The confidence of uncertainty is defined a priori and is often dependent on the number of data available.

Our proposed definition using the 95% CI criterion applied to the 251 V1 neurons analyzed in this study suggests that V1 RFs may be larger than currently appreciated. Presumably if we increase the threshold of confidence level (say, from 95% to 99%), the estimated RF size will become smaller, and the gap between these two criteria might become smaller. However, with the same level of confidence but with more spiking data collected from more experimental trials, the estimated RF size may also become smaller (due to reduced confidence bounds or reduced uncertainty for the estimated parameters).

4.3 Regularization. Maximum likelihood estimation is known to subject to overfitting by using an undesirable or unnecessary complex model. Due to the small number of trials being collected, we have not extensively investigated the regularization option for all the methods (WVK and PPGLM). Determining optimal regularization parameters and regularization schemes (e.g., L_1 norm or L_2 norm) is a nontrivial estimation problem in the presence of small sample size (Hastie, Tibshirani, & Friedman, 2009). Based on the goodness-of-fit assessment using the KS plot and cross-validation, we believe that the overfitting issue is not a concern for GLM_{hist} .

4.4 Experimental Design. Finally, our paradigm suggests an approach to determining the optimal design of moving bar experiments to maximize the accuracy of RF reconstruction. Given a fixed number of projections, the signal bandwidth that can be faithfully estimated is bounded above (Natterer, 2001). For the smoothing spline model, the number of control points per direction is bounded above by the number of moving bar projections multiplied by the time interval duration per direction. Therefore, in order to improve the estimation resolution of the RFs, it is necessary to increase the number of moving bar projections to meet the sampling criteria. Given a fixed number of trials, an experimental design can be chosen that balances the trade-off between accurate RF reconstruction and the number of directions and the number of trials per direction required to achieve it (see Figures 9 and 10).

4.5 Future directions. Our findings suggest several directions for future investigations. We can directly compare V1 RF properties (size and orientation tuning) obtained using different stimulus paradigms and examine the plasticity or sensitivity of V1 RF to distinct visual stimuli (Pettet & Gilbert, 1992; DeAngelis et al., 1995; Wandell & Smirnakis, 2009). Since our moving-bar tomographic reconstruction framework is not limited by the assumption of a linear relationship between the visual stimuli and their neural responses, it can also be used to examine the nonlinear stimulus-response relationship of neurons in other visual areas, such as V2. Alternatively, the principle can be extended to estimate the spatiotemporal RF profile (Rust, Schwartz, Movshon, & Simoncelli, 2005). In addition, instead of using filtered backprojection, alternative computational approaches for tomographic reconstruction (Herman, 2009; Mao, Fahimian, Osher, & Miao, 2010) can be used to improve the reconstruction accuracy and reduce the minimum number of moving bar directions required. Finally, our current RF model is rather simplified and limited by several assumptions. It would be interesting to explore the possibility of estimating the spatiotemporal visual RF (Theunissen et al., 2001; Stanley, 2002; Ringach, 2004; Nishimoto et al., 2005) using the same or a modified experimental paradigm. These studies will be topics of future reports.

4.6 Limitations. Finally, there are some limitations in using the moving bar stimuli paradigm. Considering a 0.2° bar width, the size of the bar width might be comparable to the full RF of some V1 cells; therefore, it might be too coarse to reconstruct the fine internal structure of RF. In addition, at the eccentricity of 2° to 5° , the fixation noise is comparable to the RF size, which might cause blurring in the reconstruction. Despite these inherent limitations caused by the original moving bar experiments, we still believe that our proposed analysis approach provides an alternative computational solution for mapping visual RFs using moving bar stimuli.

In addition, in this article, we estimate only the orientation tuning, not the directional tuning property of the V1 neurons. If V1 neurons are directionally selective (de Valois, Yund, & Hepler, 1982; Livingstone, 1998; Livingstone & Conway, 2003), then our approach would yield some biases in the RF estimate. This is due to the fact that the Radon transform is defined in the symmetric plane with orientation coverage of $[0, \pi)$ (Kak & Slanney, 1988; see also appendix A); therefore, the orientation selectivity is implicitly assumed. Despite this simplified assumption (in addition to the assumption of separability of spatial and orientation tunings), our approach still provides a useful tool to assess the visual RF properties (complementary to the STA and WVK approaches).

Since the moving bar experiment remains a classic paradigm for mapping visual RFs in awake monkeys, we believe that our proposed approach and the investigation of experimental design would provide valuable information for neuroscientists. This can also be used as a complementary

approach to compare the RF estimate obtained from the other experimental paradigms based on either random white noise stimuli or m-sequences.

Appendix A: Continuous Radon Transform

The Radon transform (Radon, 1917) is a method that has been widely used in various inverse reconstruction problems, such as computerized tomography, nuclear magnetic resonance, optics, astronomy, and geophysics (Kak & Slanney, 1988; Deans, 1993). The common task of these problems is to reconstruct a 2D or 3D function based on its lower-dimensional (1D or 2D) projections, which are treated as samples of the Radon transform.

Let $f(x, y)$ be a continuous, band-limited function defined on a 2D Euclidean space $[x, y] \in \mathbb{R}^2$. The Radon transform of f along a direction $\theta \in [0, 2\pi]$, denoted as $\mathcal{R}(f)[\theta]$, is defined by a set of line integrals along the projection direction θ at a distance t of the origin. Mathematically, the line integral is written as

$$\mathcal{R}(f)[t; \theta] = \int_{-\infty}^{\infty} f(t, s) ds, \quad (\text{A.1})$$

where (t, s) are defined by the rotated version of (x, y) as follows:

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (\text{A.2})$$

Here, $t = x \cos \theta + y \sin \theta$ defines a line in the 2D plane, which is specified by an angle θ and a distance t from the origin. Therefore, the Radon transform $\mathcal{R}(f)[\theta]$ is a 1D line projection of $f(x, y)$ for a given direction θ . The underlying mathematical theory for tomographic reconstruction is the projection-slice theorem. Let \mathcal{F}_1 and \mathcal{F}_2 denote the 1D and 2D Fourier transforms, respectively. The 2D Fourier transform of $f(x, y)$ is written as

$$\mathcal{F}_2(u, v)\{f(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i(ux+vy)} dx dy, \quad (\text{A.3})$$

where $i = \sqrt{-1}$, and $\{u, v\}$ are the spatial frequencies. Let $u = w \cos \theta$ and $v = w \sin \theta$, which convert the Cartesian coordinate system (u, v) in the frequency domain to a polar coordinate system (w, θ) ; we then obtain

$$\mathcal{F}_2(w, \theta)\{f(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i w(x \cos \theta + y \sin \theta)} dx dy. \quad (\text{A.4})$$

Because of $t = x \cos \theta + y \sin \theta$ (see equation A.2), it follows that

$$\begin{aligned}\mathcal{F}_2(w, \theta)\{f(x, y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, s) e^{-2\pi i w t} \mathrm{d}s \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(t, s) \mathrm{d}s \right] e^{-2\pi i w t} \mathrm{d}t.\end{aligned}\quad (\text{A.5})$$

From equation A.1, equation A.5 is further written as

$$\begin{aligned}\mathcal{F}_2(w, \theta)\{f(x, y)\} &= \int_{-\infty}^{\infty} \mathcal{R}(f)[t; \theta] e^{-2\pi i w t} \mathrm{d}t \\ &= \mathcal{F}_1\{\mathcal{R}\{f\}\}.\end{aligned}\quad (\text{A.6})$$

That is, the 2D Fourier transform of a function $f(x, y)$ is identical to the 1D Fourier transform of the Radon transform $\mathcal{R}\{f\}$. Note that equation A.6 is defined in the polar coordinate system (w, θ) of the frequency domain. To reconstruct $f(x, y)$, we apply the inverse 2D Fourier transform to equation A.6, which yields

$$\begin{aligned}f(x, y) &= \mathcal{F}_2^{-1}\{\mathcal{F}_1\{\mathcal{R}(f)\}\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_2(u, v) e^{2\pi i(ux+vy)} \mathrm{d}u \mathrm{d}v.\end{aligned}\quad (\text{A.7})$$

Using the identity $u = w \cos \theta$, $v = w \sin \theta$, and $\mathrm{d}u \mathrm{d}v = w \mathrm{d}w \mathrm{d}\theta$, equation A.7 is rewritten as

$$f(x, y) = \int_0^{2\pi} \int_0^{\infty} \mathcal{F}_2(w, \theta) w e^{2\pi i w(x \cos \theta + y \sin \theta)} \mathrm{d}w \mathrm{d}\theta. \quad (\text{A.8})$$

In light of the Fourier transform property, $\mathcal{F}_2(w, \theta + \pi) = \mathcal{F}_2(-w, \theta)$, equation A.8 is further written as

$$\begin{aligned}f(x, y) &= \int_0^{\pi} \int_{-\infty}^{\infty} \mathcal{F}_2(w, \theta) |w| e^{2\pi i w(x \cos \theta + y \sin \theta)} \mathrm{d}w \mathrm{d}\theta \\ &\stackrel{\text{def}}{=} \int_0^{\pi} \mathcal{Q}_{\theta}(t) \mathrm{d}\theta,\end{aligned}\quad (\text{A.9})$$

where $\mathcal{F}_2(w, \theta) = \mathcal{F}_1\{\mathcal{R}(f)[\theta]\}$, and

$$\mathcal{Q}_{\theta}(t) = \int_{-\infty}^{\infty} \mathcal{F}_1\{\mathcal{R}(f)[t; \theta]\} |w| e^{2\pi i w t} \mathrm{d}w, \quad (\text{A.10})$$

in which the frequency response $|w|$ defines a so-called Ramp filter. Hence, the reconstruction is viewed as a filtered backprojection process: a filtering operation (see equation A.10) combined with a backprojection (see equation A.9).

A.1 Numerical Issues of Filtered Backprojection. In terms of discrete implementation of the inverse Radon transform, it is worth pointing out several additional important numerical issues that are often encountered in practice (some of which might not be crucial in the computerized tomography problem but are pivotal in the RF mapping problem):

- *Ramp filter.* The original Ramp filtering is defined in continuous time. Due to discrete implementation, the exact zero frequency filtering cannot be done, so some very low-frequency components might be cancelled out. One of the solutions to alleviate this problem is to implement the Ramp filter in the time domain instead of in the frequency domain, namely, to zero-pad the filter as well as the Radon transform $\mathcal{R}(f)[t; \theta_j]$ (Kak & Slanney, 1988), followed by FFT. This is aimed at avoiding the aliasing in circular convolution and replacing the circular convolution with a nonperiodic convolution.
- *DC leakage.* If the 2D RF profile has a nonzero baseline, the DC leakage phenomenon will occur due to the discrete Fourier transform in $\mathcal{F}_1\{\mathcal{R}(f)[t; \theta_j]\}$. This issue will become more severe because of the finite projection effect (the fewer the projections, the more severe is the problem). To tackle this issue, we came up with a practical solution. Because the inverse Radon transform involves only linear operations, the linear superposition principle will hold. Therefore, we can first estimate the DC component c (the global mean value) of the reconstructed RF image and apply the Radon transform to a constant DC image (with the same image size), from which we obtain $\mathcal{R}(c)$. Before applying the inverse Radon transform, we subtract the original Radon transform by $\mathcal{R}(c)$, and finally, we apply the inverse Radon transform to $\mathcal{R}(f)[t; \theta_j] - \mathcal{R}(c)$, from which $\hat{f}(x, y)$ is obtained, and we estimate the RF by adding back the constant $f(x, y) = \hat{f}(x, y) + c$ for every pixel (x, y) . In practice, this trick is quite effective in reducing the DC leakage effect.
- *Boundary effect.* In the discrete inverse Radon transform, because of finite approximation and interpolation, the boundary values in the reconstructed RF images are less reliable (typically overestimated). In practice, we may simply ignore the estimated boundary values and display only the central image.

Appendix B: STA and WVK for RF Estimation

A visual RF can be viewed as a linear transfer function that maps the visual stimulus input to a neuronal response output. The RF characterizes the basic firing property of a neuron, either excitatory or inhibitory. In the experimental setup of mapping a 2D visual RF, we assume that the RF space has a size of $\ell \times \ell$ pixels.

Let $j = 1, \dots, J$ denote the number of bar projections. In the j th direction and at the k th time bin, let $S_{j,k} \in \mathbb{R}^{\ell^2}$ be a vectorized visual stimulus field input (represented as a $1 \times \ell^2$ row vector), and let $n_{j,k}$ denote the corresponding observed spike count (neural response output). Let $X_j = [S_{j,1}^\top, \dots, S_{j,K}^\top]^\top \in \mathbb{R}^{K \times \ell^2}$ denote a concatenated matrix that consists of K visual stimuli observed from K time instants at the j th direction (where $^\top$ denotes the transpose operation); let $n_j = [n_{j,1}, \dots, n_{j,K}]^\top \in \mathbb{R}^{K \times 1}$ denote a column vector with the concatenated neural response output at the j th direction. We further construct an ensemble stimulus matrix $X \in \mathbb{R}^{KJ \times \ell^2}$ that combines all J directions as well as an ensemble response vector $n \in \mathbb{R}^{KJ \times 1}$, such that

$$X = [X_1^\top \dots, X_J^\top]^\top,$$

$$n = [n_1^\top, \dots, n_J^\top]^\top.$$

The linear neural encoder relates the stimulus input with the neural response output via a linear regressor ξ ,

$$n = X\xi, \tag{B.1}$$

where $\xi \in \mathbb{R}^{\ell^2 \times 1}$ is a parameter vector to be estimated. Upon rematrization of ξ into an $\ell \times \ell$ matrix, we obtain the 2D RF estimate projected onto an $\ell \times \ell$ pixel space. Equation B.1 defines a linear representation between the input X and output n , and the parameter vector ξ serves as a linear transfer function.

To find a solution to linear equation B.1, we can use the least-squares method, which yields the WVK (using only linear kernels) solution,⁵

$$\xi_{\text{WVK}} = (X^\top X)^{-1} X^\top n. \tag{B.2}$$

The variance of the estimate ξ_{WVK} is

$$\text{Var}[\xi_{\text{WVK}}] = \hat{\sigma}^2 (X^\top X)^{-1}, \tag{B.3}$$

⁵This is obtained by solving the Yule-Walker equation. Multiplying X^\top to both sides of equation B.1, we approximate the autocorrelation and cross-correlation as $X^\top X$ and $X^\top n$, respectively, and it follows that $(X^\top X)\xi = X^\top n$, and the linear least-squares estimate yields the optimal Wiener solution.

where $\hat{\sigma}^2 = \frac{1}{KJ - \ell^2} \|\mathbf{n} - \mathbf{X}\boldsymbol{\xi}_{\text{WVK}}\|^2 = \frac{1}{KJ - \ell^2} (\mathbf{n} - \mathbf{X}\boldsymbol{\xi}_{\text{WVK}})^\top (\mathbf{n} - \mathbf{X}\boldsymbol{\xi}_{\text{WVK}})$ (the denominator $KJ - \ell^2$ denotes the statistical degrees of freedom).

STA can be viewed as a reduced form of the first-order WVK method. Let N denote the total number of spikes, and let $\mathbf{D}(N)$ be an $\ell^2 \times \ell^2$ diagonal matrix with N in all diagonal entries. Then the STA estimate is a special case of equation B.2 by letting $\mathbf{X}^\top \mathbf{X} = \mathbf{D}(N)$,

$$\boldsymbol{\xi}_{\text{STA}} = \mathbf{D}(N)^{-1} \mathbf{X}^\top \mathbf{n} \equiv c \mathbf{X}^\top \mathbf{n}, \quad (\text{B.4})$$

where $c = 1/N$ denotes a rate constant, which multiplies the vector $\mathbf{X}^\top \mathbf{n}$ to yield the STA estimate. That is, STA computes the normalized spike counts across time related to the stimulus; hence, it can be interpreted as coding a preferred stimulus of the neuron. In the so-called whitened STA or rotated STA (RSTA) (Paninski, 2003), the regressor vector is weighted by a full matrix rather than a constant: $\boldsymbol{\xi}_{\text{RSTA}} = \boldsymbol{\Sigma}_X^{-1} \boldsymbol{\xi}_{\text{STA}}$, where $\boldsymbol{\Sigma}_X$ denotes the sample covariance matrix computed from the stimulus samples. By simple linear algebra, we can see that the RSTA estimate is a scaled version of the WVK solution.

Next, in the context of moving bar experiments, we would like to examine the STA and WVK estimates (given data collected from all moving bar directions) in relation to their individual estimates given data collected in each direction. Put another way, can the STA or WVK estimate be composed into a linear sum of J estimates computed separately from each moving bar direction?

In the case of STA, let $N_j = \sum_{k=1}^K n_{j,k}$ denote the number of total spike counts in the j th moving bar direction, and let $N = \sum_{j=1}^J N_j$ and $c_j = 1/N_j$. We rewrite the constant c in equation B.4 as

$$c = \frac{1}{\sum_{j=1}^J N_j} = \frac{1}{\sum_{j=1}^J c_j^{-1}} = \frac{1}{J} \frac{J}{\sum_{j=1}^J c_j^{-1}}. \quad (\text{B.5})$$

Hence, c is equal to $1/J$ fraction of the harmonic mean of c_j obtained from all J directions. By decomposing the estimate $\boldsymbol{\xi}_{\text{STA}}$ in terms of individual estimates computed from each moving bar direction, equation B.4 can be rewritten as

$$\begin{aligned} \boldsymbol{\xi}_{\text{STA}} &= c \sum_{j=1}^J \mathbf{X}_j^\top \mathbf{n}_j = c \sum_{j=1}^J \boldsymbol{\xi}_j \\ &\neq \frac{1}{J} \sum_{j=1}^J c_j \boldsymbol{\xi}_j, \end{aligned} \quad (\text{B.6})$$

where $\xi_j = \mathbf{X}_j^\top \mathbf{n}_j = \sum_{k=1}^K n_{j,k} \mathbf{S}_{j,k}$ denotes the individual estimate from the j th direction. Since the harmonic mean and the arithmetic mean are generally different in value, we know from equation B.4 that $c \neq \frac{1}{J} \sum_{j=1}^J c_j$. However, in the special case when N_j or c_j is identical across all directions (i.e., the neuron firing is invariant to all orientations), the STA estimate is purely a linear average of the individual STA estimates from each direction.

Unlike STA, WVK generally cannot be decomposed into a linear combination of the individual estimates from each direction:

$$\begin{aligned} \xi_{\text{WVK}} &= \left(\sum_{j=1}^J \mathbf{X}_j^\top \mathbf{X}_j \right)^{-1} \left(\sum_{j=1}^J \mathbf{X}_j^\top \mathbf{n}_j \right) \\ &= \sum_{j=1}^J \left(\sum_{j=1}^J \mathbf{X}_j^\top \mathbf{X}_j \right)^{-1} \mathbf{X}_j^\top \mathbf{n}_j \\ &\neq \frac{1}{J} \sum_{j=1}^J (\mathbf{X}_j^\top \mathbf{X}_j)^{-1} \mathbf{X}_j^\top \mathbf{n}_j. \end{aligned} \quad (\text{B.7})$$

From simple linear algebra, we know in general that the last two equations are not equal in value, except for a special case when $\mathbf{X}_j^\top \mathbf{X}_j$ are identical for all j . In the moving bar experiment, if the bar length is greater than the diameter or maximum span of the visual field, this special condition could be satisfied since the area coverage of the moving bar, $\mathbf{X}_j = [\mathbf{S}_{j,1}^\top, \dots, \mathbf{S}_{j,K}^\top]^\top$, is exactly the same at each direction in a circular-shape visual field (due to circular symmetry).

Note that in equation B.7, the matrix $\mathbf{X}_j^\top \mathbf{X}_j$ has to be of full rank for all j th directions, which typically requires a large number of K for each direction. Because of the least-squares fit, the elements of $\mathbf{X}_j \xi_{\text{WVK}}$ are not always nonnegative (see, e.g., Figure 5a, right-most column), and the least-squares estimate might be sensitive to noise or outliers in \mathbf{X}_j , as reflected in the worst performance in goodness of fit (see Table 1).

Appendix C: Data Recordings and Visual Stimulus Setup

The experimental procedures were approved by the National Committee on Animal Welfare in compliance with the guidelines of the European Community for the care and use of laboratory animals. Two adult female rhesus monkeys (*Macaca mulatta*) were used in this study. Neuronal spiking activity was recorded from awake and head-fixed monkeys in the opercular region of V1 (2–5° of eccentricity) and, on some occasions, from the superior bank of the calcarine sulcus (8–12° of eccentricity). The receptive fields were

first visualized using a method that was developed by Mario Fiorani in the Biophysics Institute Carlos Chagas Filho at the Federal University of Rio de Janeiro, Brazil (Fiorani et al., 2001).

Each trial started with an appearance of a red fixation point (0.15° ; luminance: $10 \text{ cd}\cdot\text{m}^{-2}$), and the monkeys were trained to press a lever within 700 ms and to maintain their gazes within a virtual window (window size: 1°) centered on the fixation point. The task ended by randomly changing the color of the fixation point from red to green between 2500 and 4000 ms after the trial onset. To obtain a juice reward, monkeys had to release the lever within 500 ms after the color change in the fixation point. Trials were discarded whenever fixation was interrupted. The monkeys' eye positions were tracked continuously by a search coil system (DNI and Crist Instruments) with a 500 Hz sampling rate.

Visual stimuli were generated as sequences of computer-generated bitmap images with an interface developed in LabVIEW (National Instruments) and were presented as movies (frame rate: 100 frames/s) using a standard graphic board (GeForce 6600-series, NVIDIA) controlled by ActiveSTIM (www.activestim.com). The CRT monitor (CM813ET, Hitachi) was gamma corrected and spanned a visual angle of $36^\circ \times 28^\circ$. The visual stimulus was a high-contrast light bar (luminance: $50 \text{ cd}\cdot\text{m}^{-2}$; bar width: 0.2°) that was moved with a constant velocity ($14.9^\circ/\text{s}$) in the visual field ($21.8^\circ \times 21.8^\circ$). The bar stimulus was presented in random order in 16 different orientations separately and evenly by 22.5° . Each trial lasted 2 seconds (during which the bar traveled inside the visual field), where $t = 0$ and $T = 2 \text{ s}$, respectively, corresponded to the moment when the bar's position overlaid two corners of the visual field in the diagonal and $t = 1$ corresponded to the center of the field (see Figure 1b).

Quartz-insulated tungsten-platinum electrodes ($\sim 80 \mu\text{m}$ diameter, $0.3\text{--}1.0 \text{ M}\Omega$ impedance; Thomas Recording) were used to record the extracellular neural activity from three to five sites in both superficial and deep layers of the striate cortex (digitally bandpass filtered, $0.7\text{--}6.0 \text{ kHz}$; Plexon). Neuronal spikes were detected by amplitude thresholding, which was set interactively based on online visualization of the spike waveforms. Spike events and corresponding waveforms were sampled at 32 kHz (spike waveform length: 1.2 ms). Single-unit activity was separated using an interactive custom-designed spike-sorting software. Experimental trials with artifacts were rejected during which either monkey did not maintain fixation or showed no response or incorrect behavior. Upon trial rejection, the complete data sets were balanced for all conditions such that all recorded V1 neurons contained exactly nine trials per direction.

Acknowledgments

We thank two anonymous reviewers for valuable comments. This work was supported by the U.S. National Institutes of Health, grant DP1-OD003646

(to E.N.B.), and the European Union, FP7-1CT, grant 240763 (to G.P.). Some preliminary results were presented in COSYNE'09, Salt Lake City, Utah (Pipa, Chen, Neuenschwander, Lima, & Brown, 2009).

References

- Brown, E. N., Barbieri, R., Eden, U. T., & Frank, L. M. (2003). Likelihood methods for neural data analysis. In J. Feng (Ed.), *Computational neuroscience: A comprehensive approach* (pp. 253–286). Boca Raton, FL: CRC Press.
- Brown, E. N., Barbieri, R., Ventura, V., Kass, R. E., & Frank, L. M. (2002). The time-rescaling theorem and its application to neural spike train analysis. *Neural Comput.*, 14, 325–346.
- Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nat. Neurosci.*, 7, 456–461.
- Chen, X., Han, F., Poo, M. M., & Dan, Y. (2007). Excitatory and suppressive receptive field subunits in awake monkey primary visual cortex (V1). *Proc. Natl. Acad. Sci. USA*, 104, 19120–19125.
- Crist, R. E., Li, W., & Gilbert, C. D. (2001). Learning to see: Experience and attention in primary visual cortex. *Nat. Neurosci.*, 4, 519–525.
- Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res.*, 20, 847–856.
- de Boor, C. (1978). *A practical guide to splines*. Berlin: Springer-Verlag.
- de Ruyter van Steveninck, R. R., & Bialek, W. (1988). Coding and information transfer in short spike sequences. *Proc. R Soc. Lond. B*, 234, 379–414.
- De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vis. Res.*, 22, 545–559.
- De Valois, R. L., Yund, W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vis. Res.*, 22, 531–544.
- DeAngelis, G. C., Anzai, A., Ohzawa, I., & Freeman, R. D. (1995). Receptive field structure in the visual cortex: Does selective stimulation induce plasticity? *Proc. Natl. Acad. Sci. USA*, 92, 9682–9686.
- Deans, S. R. (1993). *The radon transform and some of its applications*. New York: Dover.
- Fiorani, M., Azzi, J. C., & Gattass, R. (2001). Primate V1 has a visuotopic rather than a retinotopic organization: The topographic organization of the blind spot (BS) in V1. *Soc. Neurosci.* 27:619.46 (Abstract).
- Gabbiani, F. (1996). Coding of time varying signals in spike trains of linear and half-wave rectifying neurons. *Network*, 7, 61–85.
- Haslinger, R., Pipa, G., & Brown, E. N. (2010). Discrete time rescaling theorem: Determining goodness of fit for discrete time statistical models of neural spiking. *Neural Comput.*, 22, 2477–2506.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Berlin: Springer.
- Herman, G. T. (2009). *Fundamentals of computerized tomography: Image reconstruction from projection* (2nd ed.). Berlin: Springer.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.*, 148, 574–591.

- Jacobs, A. L., Fridman, G., Douglas, R. M., Alam, N. M., Latham, P. E., Prusky, G. T., et al. (2009). Ruling out and ruling in neural codes. *Proc. Natl. Acad. Sci. USA*, 106, 5936–5941.
- Jones, J. P., & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.*, 58, 1187–1211.
- Kak, A. C., & Slaney, M. (1988). *Principles of computerized tomographic imaging*. New York: IEEE Press.
- Kass, R. E., Ventura, V., & Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *J. Neurophysiol.*, 94, 8–25.
- Leventhal, A. G., Thompson, K. G., Liu, D., Zhou, Y., & Ault, S. J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *J. Neurosci.*, 15, 1808–1818.
- Livingstone, M. S. (1998). Mechanisms of direction selectivity in macaque V1. *Neuron*, 20(3), 509–526.
- Livingstone, M. S., & Conway, B. R. (2003). Substructure of direction-selective receptive fields in macaque V1. *J. Neurophysiol.* 89(5), 2743–2759.
- Mao, Y., Fahimian, B. P., Osher, S. J., & Miao, J. (2010). Development and optimization of regularized tomographic reconstruction algorithms utilizing equally-sloped tomography. *IEEE Trans Image Process*, 19, 1259–1268.
- Mardia, K. V. (1972). *Statistics of directional data*. Orlando, FL: Academic Press.
- McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. London: Chapman & Hall.
- Movellan, J. R. (2008). *Tutorial on Gabor filter* (Tech. Rep.). San Diego: University of California, San Diego.
- Movshon, J. A., Thomson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.*, 283, 53–77.
- Natterer, F. (2001). *The mathematics of computerized tomography*. Philadelphia: SIAM Press.
- Nienborg, H., Bridge, H., Parker, A. J., & Cumming, B. G. (2004). Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *J. Neurosci.*, 24, 2065–2076.
- Nishimoto, S., Arai, M., & Ohzawa, I. (2005). Accuracy of subspace mapping of spatiotemporal frequency domain visual receptive fields. *J. Neurophysiol.*, 93, 3524–3536.
- Paninski, L. (2003). Convergence properties of some spike-triggered analysis techniques. *Network*, 14, 437–464.
- Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network*, 15, 243–262.
- Park, M., & Pillow, J. W. (2011). Receptive field inference with localized priors. *PLoS Comput. Biol.*, 7(10), e1002219.
- Pawitan, Y. (2001). *In all likelihood: Statistical modelling and inference using likelihood*. New York: Oxford University Press.
- Pettet, M. W., & Gilbert, C. D. (1992). Dynamic changes in receptive-field size in cat primary visual cortex. *Proc. Natl. Acad. Sci. USA*, 89, 8366–8370.
- Pipa, G., Chen, Z., Neuenschwander, S., Lima, B., & Brown, E. N. (2009). Efficient spike encoding for mapping visual receptive fields. Oral presentation in *Computational and Systems Neuroscience (COSYNE'09)*. Abstract in *Frontier in Neuroscience*.

- Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte Sächsische Akademie der Wissenschaften. Leipzig, Math.-Phys. Kl.*, 69, 262–267.
- Reid, R. C., Victor, J. D., & Shapley, R. M. (1997). The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Vis. Neurosci.*, 14, 1015–1027.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- Ringach, D. L. (2004). Mapping receptive fields in primary visual cortex. *J. Physiol.*, 558, 717–728.
- Ringach, D. L., & Shapley, R. M. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, 28, 147–166.
- Ringach, D. L., Shapley, R. M., & Hawken, M. J. (2002). Orientation selectivity in macaque V1: Diversity and laminar dependence. *J. Neurosci.*, 22, 5639–5651.
- Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res.*, 5, 583–601.
- Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque V1 receptive fields. *Neuron*, 46, 945–956.
- Shimokawa, T., & Shinomoto, S. (2009). Estimating instantaneous irregularity of neuronal firing. *Neural Computation*, 21, 1931–1951.
- Stanley, G. B. (2002). Adaptive spatiotemporal receptive field estimation in the visual pathway. *Neural Comput.*, 14, 2925–2946.
- Sun, M., & Bonds, A. B. (1994). Two-dimensional receptive-field organization in striate cortical neurons of the cat. *Vis. Neurosci.*, 11, 703–720.
- Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E., & Gallant, J. L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network*, 12, 289–316.
- Truccolo, W., Eden, U. T., Fellow, M., Donoghue, J. P., & Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and covariate effects. *J. Neurophysiol.*, 93, 1074–1089.
- Truccolo, W., Hochberg, L. R., & Donoghue, J. P. (2010). Collective dynamics in human and monkey sensorimotor cortex: Predicting single neuron spikes. *Nat. Neurosci.*, 13, 105–111.
- Wandell, B. A., & Smirnakis, S. M. (2009). Plasticity and stability of visual field maps in adult primary visual cortex. *Nat. Rev. Neurosci.*, 10, 873–884.

Copyright of Neural Computation is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.